SOE-1
B-096E
C.2.

Summary Report on the Quality Control of the Data Capture for the 1981 Census of Agriculture

Quality Control Section
I.A.S.M.D.
January 1982

1981 Census of Agriculture

Quality Control of the Data Capture

A. INTRODUCTION

This report summarizes the application of quality control on the data capture operation for the 1981 Census of Agriculture.

The purpose of this project was to attempt to achieve a pre-specified quality level or better at the time of data capture, and simultaneously minimize the time and cost of verification by implementing a Sample Verification methodology. The Average Outgoing Quality Limit (AOQL) was set at 4.0%, which simply means that at least 96% of all the records processed would be free of data entry errors, after the application of the Quality Control process.

B. METHOD OF QUALITY CONTROL

1. INSTRUCTIONS

-The detailed keying instructions were prepared by the Census Survey Methods Sub-Division, C.H.S.M.D. in conjunction with the Data Capture Section of D.P.D.

-The verification procedures were prepared by the Quality Control Section, I.A.S.M.D.

2. TRAINING OF VERIFIERS

The training of the staff for sample verification was conducted by the Q.C. Section in DPD during the evening shift (i.e. 4:00 p.m. to 7:30p.m.)

The training schedule was as follows:

TRAINING SCHEDULE FOR VERIFIERS

DATE	NO. OF VERIFIERS
July 13	4
July 14	4
July 20	4
July 21	4
Aug. 4	4
Aug. 6	3 + 1 Lead Operator
Aug. 11	4 + 1 Supervisor
Aug. 13	(English) > Session with all operators to
Aug. 14	(French) discuss problems
Aug. 18	4 + 1 Supervisor
Aug. 25	4 + 1 Lead Operator
Aug. 27	4 + 1 Lead Operator
Sept. 1	Session with all verifiers to discuss results & problems
Sept. 8	Session with all verifiers to discuss results & problems
Sept.17	Session with all verifiers to discuss results & problems
Sept.29	Session with all verifiers to discuss results & problems
Nov. 5	Session with all verifiers to discuss results & problems

Ten training sessions of approximately 3 hours each were conducted. Each session consisted of oral instruction, discussions and a written test.

A total of 36 verifiers, 3 lead operators and 2 supervisors were trained. Generally the training sessions went well except for three verifiers who showed difficulties and had to attend a second session.

Seven additional review sessions were held with all operators to discuss any problems encountered and to ensure that the procedures were being fully understood.

3. QUALIFICATION OF KEYERS

Initially, all keying operators were placed on a 100% verification plan until they processed a sufficient number of records to reliably estimate their Error Rates.

Several problems were encountered at the start of processing and consequently the qualification of keyers onto sample verification was much slower than expected. In an effort to qualify them as quickly as possible their Error Rates were calculated daily. Whenever a keyers error rate stabilized at 8% or lower, she/he was assigned a Sampling Plan accordingly. The following is a schedule of qualifications made by processing week.

QUALIFICATION SCHEDULE FOR KEYERS

WEEK	DATE	NO. OF KEYERS QUALIFIED	NO. OF QUAL. KEYERS PROCESSING	WEEK	DATE	NO. OF KEYERS QUALIFIED	NO. OF QUAL. KEYERS PROCESSING
1	Aug. 5	2		6	Sept. 4	2	
1	Aug. 6	1		6	Sept. 8	2	33
2	-	-	2	7	Sept. 14	1	33
3	Aug. 17	1		8	Sept. 15	5	
3	Aug. 20	8		8	Sept. 22	1	37
3	Aug. 21	1	4	9	Sept. 28	1	35
4	Aug. 24	2		10	Oct. 2	1	38
4	Aug. 25	5		11	Oct. 9	3	42
4	Aug. 26	2	19	12			40
5	Aug. 31	4		13			40
5	Sept. 1	1		14			38
5	Sept. 2	1		15			36
5	Sept. 3	4	24	16			35
				17			33
				18			33

In all a total of 48 keyers were qualified for sample verification. The majority of initial qualifications were on Plans of 7.0% and 8.0% which were subsequently revised weekly (generally downward) according to the keyers processing abilities.

4. VERIFICATION TESTS

The implementation of sample verification usually includes some examination of each verifier's performance through verification tests. This is to ensure that the procedures are being fully understood and applied correctly. The verifiers were told that their performance would be evaluated via these tests. However, due to the many problems experienced at the start of processing (i.e. not relating to verification), the tight scheduling constraints and the length of time that would be required to perform these tests, they were never applied. The Sample Verification process was simply monitored on an overall general basis by comparing expected statistics to observed data. (See Table on page 6)

5. FEEDBACK INFORMATION

a. Operator Feedback

i) Daily

The keyers were informed every day of outstanding errors (and possilbe mis-interpretations) as well as their progress.

Feedback was also given to the verifiers on their accuracy in noting errors on the Control Card.

ii) Weekly

-Cells in Errors:
For the first 9 weeks, all cells in error were compiled and the frequency of erros by cell ID was given to each keyer.

From the 10th week on, cells in errors were given only to keyers still on 100% verification.

-Error Rates:
The keyers were also given their respective Error Rates; they were very concerned about their individual and relative performances.

-Personal Operator Information:
This printout containing information on each lot processed was given to all keyers.

b. Supervisor Feedback

i) Daily

The supervisors were contacted daily to keep them informed on operator performances and survey progress.

ii) Weekly

Every week, the supervisors were provided with a list of the keyers and their revised sampling plans.

6. WEEKLY PROCESSING OF Q.C. DATA

The completed Control Cards were collected daily with a weekly cut-off of Wednesdays. The Control Cards were coded and then on Thursday morning they were submitted to DPD for keying which was usually completed the same day. The Q.C. data was then run through our Edit and Q.C. Tabulation programs. The results were subsequently analyzed and by Friday of each week, the relevant feedback information was distributed to both the operators and the supervisors.

This weekly schedule was initiated during the 4th week of processing (Friday, Aug. 28) and maintained until the survey was completed (Friday, Nov. 27).

C. QUALITY CONTROL IMPLEMENTATION

Quality Control was implemented on August 3, however, many problems were encountered during the first month of processing which delayed the qualification of keyers onto sample verification.

In general, the operators found the data entry operation quite complex which seemed to cause them a lot of confusion and frustration. Discussion sessions were arranged with the supervisors, lead operators, verifiers and keyers, and the following problems were identified:

-Documents
- ambiguous, multiple and unreadable entries;
- confusion with accolades (}), plus (+) and minus (-) signs beside the data;
- different pencil and ink color entries and notations; etc.

-Keying Instructions
- some parts were confusing;
- generally quite lengthy with many exception conditions.

The ambiguities and confusion were causing many disagreements between the verifiers and the keyers which subsequently resulted in much higher than normal Error Rates. At this stage, these rates were recognized as being more disagreement rates rather than error rates, since in many cases, the verifier made correct entries incorrect.

The problems were discussed with all members of the project team and the following actions were taken:

-Census of Agriculture sub-division increased the amount of manual editing at Prescan, thereby clarifying many of the most difficult entries.

-Data Processing Division streamlined the keying instructions.

-Quality Control Section continued to provide daily feedback to

operators and to monitor the Error Rates closely. Also provided was additional training (i.e. via additional review sessions) to all operators in an effort to reduce the amount of keyer-verifier disagreements.

Based on these actions, the data capture operation for this survey began progressing much more smoothly with subsequent processing i.e. starting in September. Every week thereafter, the overall Verification and Error Rates decreased gradually to acceptable levels. (see table on next page)

D. RESULTS AND ANALYSIS

The following are the highlights of the quality control results which were obtained during the processing of the survey:

Total No. of Records Keyed	931,966
Total No. of Records Verified	201,960
Overall Verification Rate	41.8%
Sample Verification Rate	27.1%
Expected Sample Verification Rate.	30.3%
Average Incoming Quality (Error Rate).	6.3%
Average Outgoing Quality (AOQ)	2.8%

During the Volume Test it was predicted that the overall error rate would fall between 4% and 5%. The error rate actually measured overall operators during the processing of the survey was 6.3%. However, as has already been noted, the many unexpected problems encountered during the first month of processing contributed significantly to this higher error rate. It is important to note that after this period (i.e. after the incorporation of additional pre-scanning, training and streamlining the keying instructions) the overall error rate dropped to 5.4%. Furthermore, the error rate for all sample qualified operators averaged 4.8%.

Because of the higher than expected starting error rates, a large amount of initial 100% verification was required. This subsequently resulted in an overall verification rate of 41.8%. Again, however, this rate was only 34.3% after the first month of processing.

The sample verification rate was measured at 27.1% and this compared reasonably well to the 30.3% expected (i.e. theoretical calculation using expected probabilities of acceptance). The rejection rate was measured at 11.8%. This was a little lower than the 15.7% expected which was calculated at the observed levels of incoming quality. This probably indicated that a small amount of under-rejection occurred during the application of the quality control process.

The Average Outgoing Quality was measured at 2.8% and this, as well, compared favourably to the AOQL of 4.0%. The following table is a weekly summary of the data processed.

1981 CENSUS OF AGRICULTURE

WEEKLY Q.C. SUMMARY
AOQL = 4%

WEEK	CUT-OFF DATE	NO. OF LOTS	NO. OF REC. KEYED	NO. OF REC. VERIF. SAMPLE	NO. OF REC. VERIF. 100%	VERIFICATION RATE OVERALL	VERIFICATION RATE SAMPLE ACT.	VERIFICATION RATE SAMPLE EXP.	REJECTION RATE ACT.	REJECTION RATE EXP.	ERROR RATE (OVERALL)	AOQ OVERALL
1	Aug. 6	82	16,726	0	16,726	100.0	-	-	-	-	16.3	0.0
2	Aug. 13	169	33,995	365	33,630	100.0	-	-	-	-	15.6	0.0
3	Aug. 20	171	34,224	1,203	30,195	91.7	29.9	-	14.3	-	10.9	0.5
4	Aug. 26	213	42,726	4,305	27,080	73.9	28.8	33.6	10.1	15.9	9.3	1.4
5	Sep. 2	260	49,988	7,307	21,997	58.6	26.1	31.6	8.3	15.0	7.2	2.1
6	Sep. 7	228	45,072	8,531	12,172	45.9	25.9	29.6	7.2	12.0	6.2	2.5
7	Sep. 16	282	58,268	11,714	16,133	47.8	27.8	28.7	11.3	12.8	6.5	2.4
8	Sep. 23	336	65,799	14,199	11,996	39.8	26.4	28.2	10.6	13.1	5.7	2.7
9	Sep. 30	333	66,596	16,395	7,561	36.0	27.8	32.5	12.2	18.2	6.0	3.3
10	Oct. 7	382	75,001	17,517	7,440	34.6	27.4	29.7	12.8	15.2	5.2	3.0
11	Oct. 14	254	52,912	15,536	1,861	32.9	30.4	32.4	15.6	18.4	5.2	3.4
12	Oct. 21	353	73,341	19,923	289	27.6	27.3	28.6	12.5	14.2	4.5	3.3
13	Oct. 28	317	66,282	16,184	48	24.4	24.4	28.7	9.5	15.1	4.5	3.4
14	Nov. 4	320	66,032	16,215	49	24.5	24.5	29.1	8.8	13.4	4.6	3.4
15	Nov. 10	192	38,746	11,529	84	29.8	29.6	34.9	16.0	19.9	5.4	3.8
16	Nov. 18	225	48,897	14,507	26	29.7	29.6	32.2	15.6	19.3	5.3	3.7
17	Nov. 25	237	47,648	11,676	97	24.5	24.4	30.0	10.3	16.6	4.7	3.6
18	Dec. 2	247	49,638	13,958	206	28.5	28.2	30.4	14.0	17.0	4.8	3.4
	TOTAL	4,601	931,966	201,960	187,590	41.8 (34.3)*	27.1	30.3	11.8	15.7	6.3 (5.4)*	2.8

Footnotes to Summary Table

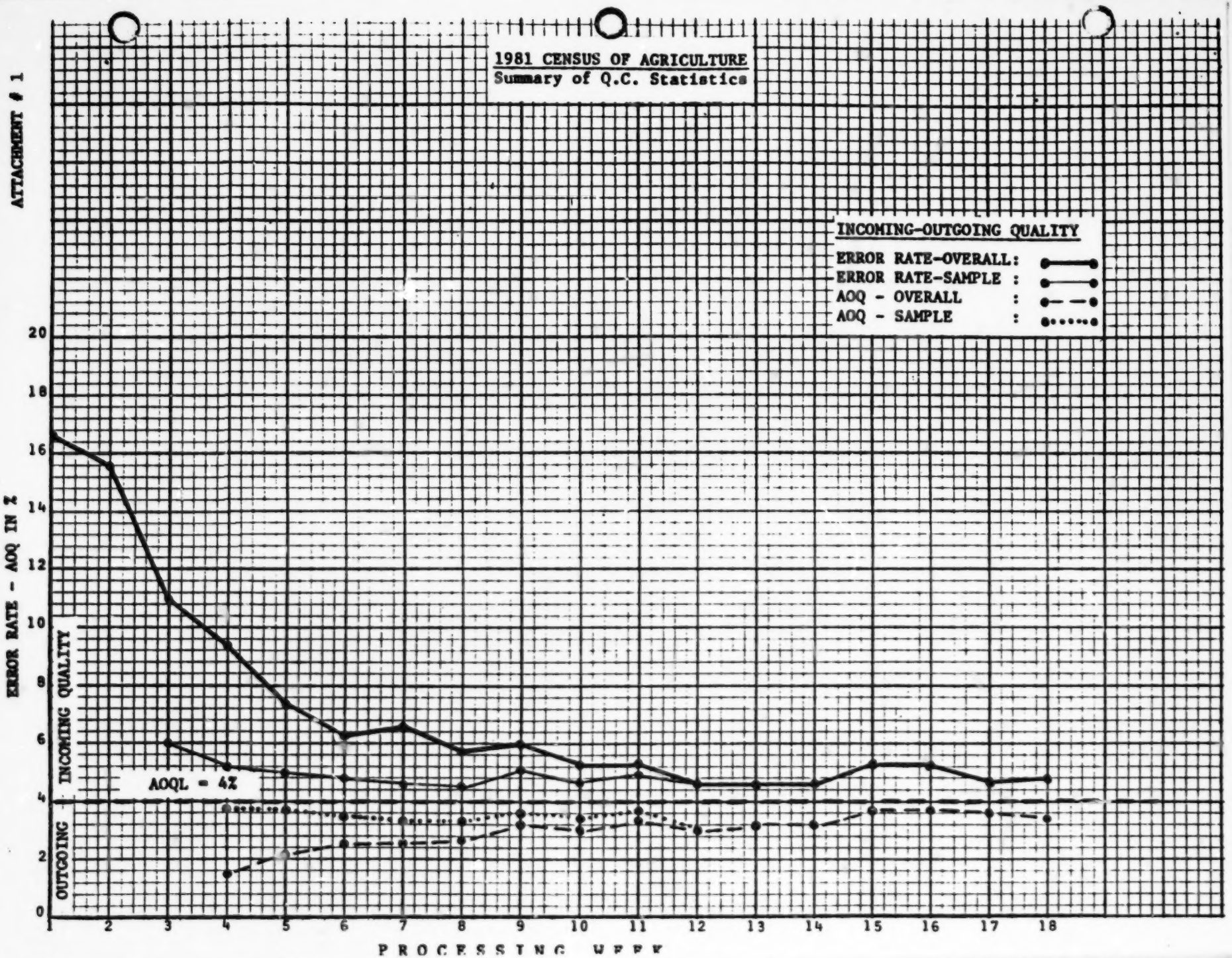
1 - Week 2: Only two lots were sample verified and both were rejected.
2 - Week 2-3: The number of lots sample verified is too small to calculate meaningful sample expected rates.
3 - Some weeks indicated that the Expected Verification and Rejection Rates were sufficiently higher than the actual rates to suspect a certain amount of under-rejection of lots. Therefore, several sessions were held with the verifiers to discuss this problem and at the same time, review the Sample Verification procedures.
4 - The starred (*) numbers in the Totals relate to rates calculated after the first month of processing.

E. ATTACHMENTS

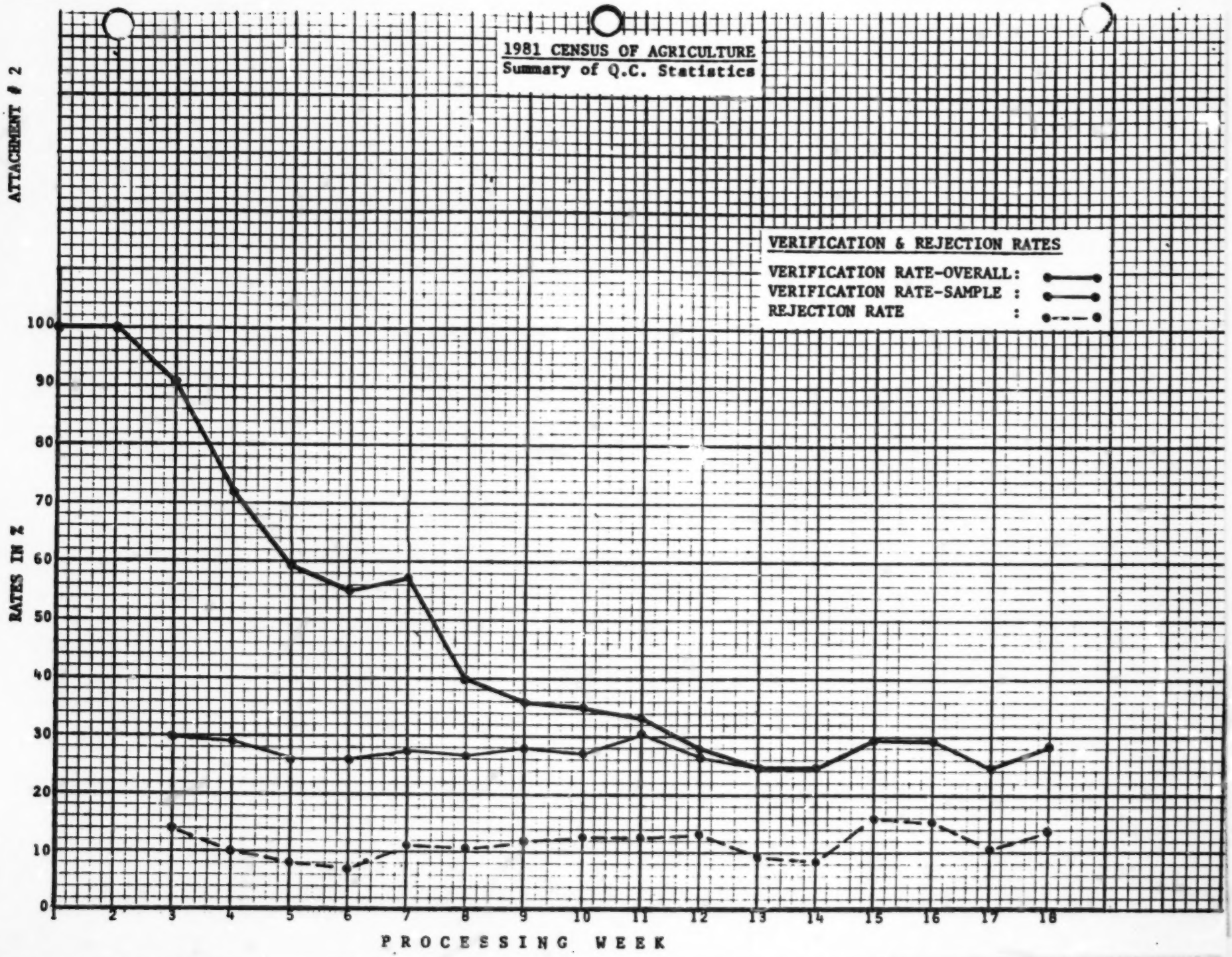
The following attachments are provided

1 - Graph of Incoming vs. Outgoing Quality by week
2 - Graph of Verification & Rejection Rates by week
3 - Cumulative Distribution of Verification Rates by week
4 - A detailed list of Cells in Error by Cell ID, their frequency of error and their distribution as a % of the total error.

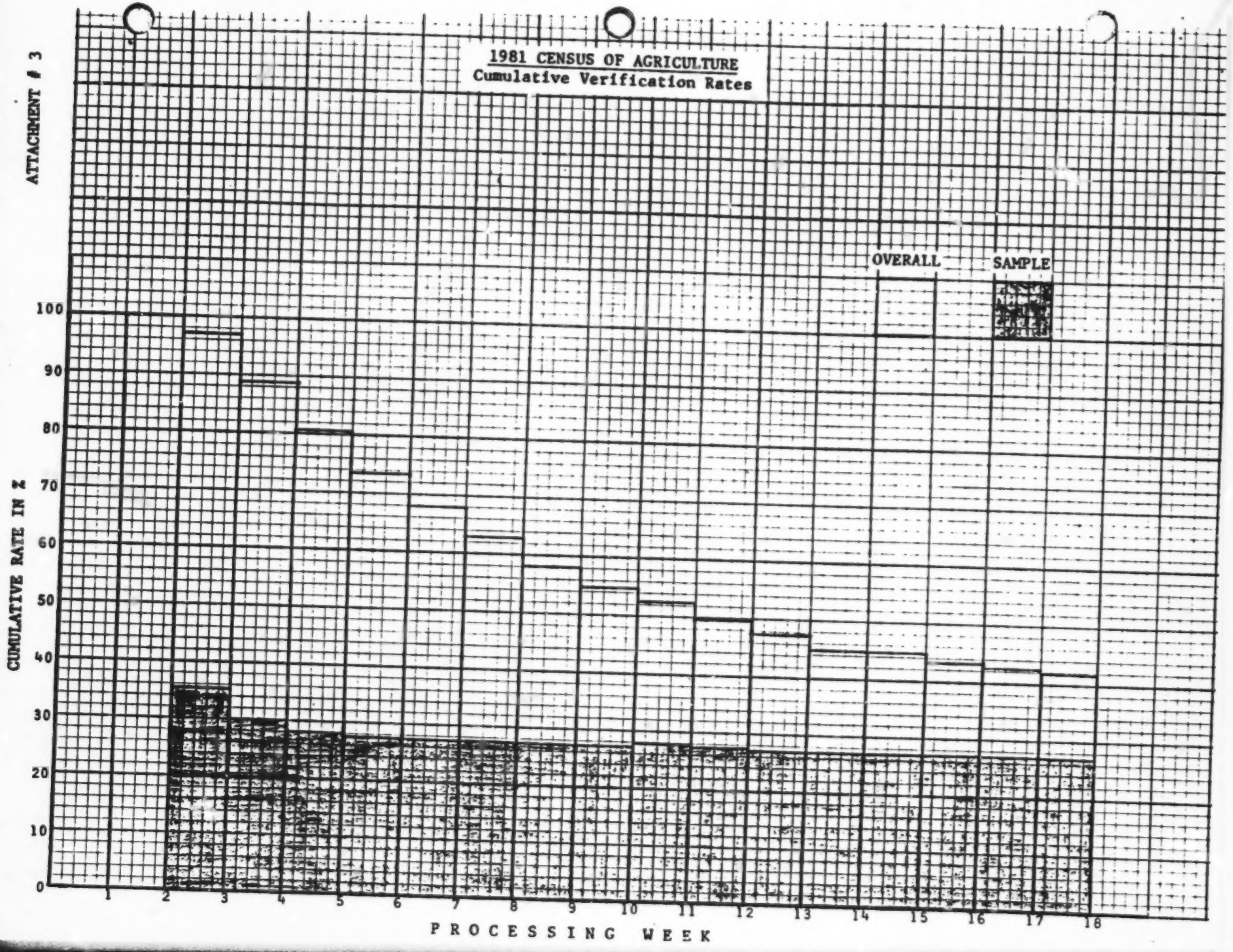
ATTACHMENT # 1



ATTACHMENT # 2



ATTACHMENT # 3



ATTACHMENT # 4

1981 CENSUS OF AGRICULTURE
CELLS IN ERROR

CELL ID	FREQ	% Dist	CELL ID	FREQ	% DIST	CELL ID	FREQ	% DIST	CELL ID	FREQ	% DIST	CELL ID	FREQ	% DIST
001	39	0.1	061	6	0.02	109	9	0.02	157	93	0.2	205	175	0.5
002	34	0.09	062	6	0.02	110	9	0.02	158	159	0.4	206	65	0.2
003	69	0.2	063	3	0.008	111	15	0.04	159	299	0.8	207	114	0.3
004	750	2.0	064	-	0	112	3	0.008	160	422	1.1	208	209	0.5
005	568	1.5	065	1	0.003	113	10	0.03	161	556	1.4	209	277	0.7
006	47	0.1	066	4	0.01	114	36	0.09	162	459	1.2	210	28	0.07
007	132	0.3	067	3	0.008	115	20	0.05	163	49	0.1	211	49	0.1
008	306	0.8	068	1	0.003	116	11	0.03	164	70	0.2	212	74	0.2
009	137	0.4	069	29	0.07	117	14	0.04	165	99	0.3	213	100	0.3
010	62	0.2	070	41	0.1	118	6	0.02	166	103	0.3	214	741	1.9
011	85	0.2	071	57	0.2	119	34	0.09	167	67	0.1	215	1061	2.8
024	46	0.1	072	225	0.6	120	8	0.02	168	122	0.3	216	111	0.3
025	174	0.5	073	15	0.04	121	49	0.1	169	32	0.08	217	229	0.6
026	189	0.5	074	43	0.1	122	97	0.2	170	23	0.06	218	463	1.2
027	128	0.3	075	68	0.2	123	75	0.2	171	27	0.07	219	169	0.4
028	2	0.005	076	10	0.03	124	7	0.02	172	151	0.4	220	285	0.7
029	814	2.1	077	37	0.1	125	748	1.9	173	45	0.1	221	665	1.7
030	872	2.3	078	64	0.2	126	440	1.1	174	39	0.1	222	642	1.7
031	770	2.0	079	14	0.04	127	250	0.7	175	2	0.005	223	679	1.8
032	421	1.1	080	35	0.09	128	695	1.8	176	4	0.01	224	519	1.4
033	1089	2.8	081	71	0.2	129	462	1.2	177	7	0.02	225	362	0.9
034	155	0.4	082	4	0.01	130	437	1.1	178	40	0.1	226	731	1.9
035	43	0.1	083	13	0.03	131	815	2.1	179	36	0.09	227	710	1.9
036	23	0.06	084	27	0.07	132	72	0.2	180	165	0.4	228	570	1.5
037	61	0.2	085	8	0.02	133	11	0.03	181	83	0.3	229	592	1.5
038	190	0.5	086	13	0.03	134	5	0.01	182	146	0.4	230	151	0.4
039	166	0.4	087	50	0.1	135	17	0.04	183	214	0.6	231	92	0.2
040	76	0.2	088	7	0.02	136	45	0.1	184	68	0.2	232	835	2.2
041	131	0.3	089	23	0.06	137	3	0.008	185	70	0.2	233	87	0.2
042	22	0.06	090	32	0.08	138	28	0.07	186	117	0.3	234	245	0.6
043	11	0.03	091	135	0.4	139	18	0.05	187	37	0.1	INS/DEL	605	1.6
044	8	0.02	092	66	0.2	140	333	0.9	188	368	1.0	QUEST NO	65	0.2
045	19	0.05	093	97	0.3	141	441	1.1	189	680	1.8	ENV NO	90	0.2
046	49	0.1	094	49	0.1	142	161	0.4	190	320	0.8	TOT	38,376	100.0
047	4	0.01	095	65	0.2	144	159	0.4	191	727	1.9			
048	28	0.07	096	31	0.08	144	92	0.2	192	331	0.9			
049	2	0.005	097	30	0.08	145	109	0.3	193	470	1.2			
050	90	0.2	098	49	0.1	146	269	0.7	194	251	0.7			
051	216	0.6	099	52	0.1	147	135	0.4	195	429	1.1			
052	260	0.7	100	35	0.09	148	190	0.4	196	101	0.3			
053	60	0.2	101	82	0.2	149	47	0.1	197	231	0.6			
054	47	0.1	102	54	0.1	150	35	0.09	198	55	0.1			
055	29	0.07	103	30	0.08	151	168	0.4	199	154	0.4			
056	19	0.05	104	44	0.1	152	165	0.4	200	113	0.3			
057	4	0.01	105	41	0.1	153	170	0.4	201	245	0.6			
058	9	0.02	106	155	0.4	154	238	0.6	202	59	0.2			
059	11	0.03	107	101	0.3	155	197	0.5	203	110	0.3			
060	-	0	108	7	0.02	156	172	0.4	204	108	0.3			

This table is a compilation of the number of times each cell was recorded in error on the Control Card.
The % Distribution column is calculated as a percentage of the total number of errors identified. Relative frequencies are not shown since the frequency of occurence for each cell is not known.

1981 CENSUS OF AGRICULTURE

WEEKLY Q.C. SUMMARY
AOQL = 4%

WEEK	CUT-OFF DATE	NO. OF LOTS	NO. OF REC. KEYED	NO. OF REC. VERIF.		VERIFICATION RATE			REJECTION RATE		ERROR RATE	AOQ OVERALL
				SAMPLE	100%	OVERALL	SAMPLE					
							ACT.	EXP.	ACT.	EXP.	(OVERALL)	
1	Aug. 6	82	16,726	0	16,726	100.0	-	-	-	-	16.3	0.0
2	Aug. 13	169	33,995	365	33,630	100.0	-	-	-	-	15.6	0.0
3	Aug. 20	171	34,224	1.203	30,195	91.7	29.9	-	14.3	-	10.9	0.5
4	Aug. 26	213	42,726	4,505	27,080	73.9	28.8	33.6	10.1	15.9	9.3	1.4
5	Sep. 2	260	49,988	7,307	21,997	58.6	26.1	31.6	8.3	15.0	7.2	2.1
6	Sep. 7	228	45,072	8,531	12,172	45.9	25.9	29.6	7.2	12.0	6.2	2.5
7	Sep. 16	282	58,268	11,714	16,133	47.8	27.8	28.7	11.3	12.8	6.5	2.4
8	Sep. 23	336	65,799	14,199	11,996	39.8	26.4	28.2	10.6	13.1	5.7	2.7
9	Sep. 30	333	66,596	16,395	7,561	36.0	27.8	32.5	12.2	18.2	6.0	3.3
10	Oct. 7	382	75,001	17,517	7,440	34.6	27.4	29.7	12.8	15.2	5.2	3.0
11	Oct. 14	254	52,912	15,536	1,861	32.9	30.4	32.4	15.6	18.4	5.2	3.4
12	Oct. 21	353	73,341	19,923	289	27.6	27.3	28.6	12.5	14.2	4.5	3.3
13	Oct. 28	317	66,282	16,184	48	24.4	24.4	28.7	9.5	15.1	4.5	3.4
14	Nov. 4	320	66,032	16,215	49	24.5	24.5	29.1	8.8	13.4	4.6	3.4
15	Nov. 10	192	38,746	11,529	84	29.8	29.6	34.9	16.0	19.9	5.4	3.8
16	Nov. 18	225	48,897	14,507	26	29.7	29.6	32.2	15.6	19.5	5.3	3.7
17	Nov. 25	237	47,648	11,676	97	24.5	24.4	30.0	10.3	16.6	4.7	3.6
18	Dec. 2	247	49,638	13,958	206	28.5	28.2	30.4	14.0	17.0	4.8	3.4
	TOTAL	4,601	931,966	201,960	187,590	41.8 (34.3)*	27.1	30.3	11.8	15.7	6.3 (5.4)*	2.8

1981 CENSUS OF AGRICULTURE

WEEKLY Q.C. SUMMARY
AOQL = 4%

WEEK	CUT-OFF DATE	NO. OF LOTS	NO. OF REC. KEYED	NO. OF REC. VERIF.		VERIFICATION RATE			REJECTION RATE		ERROR RATE	AOQ OVERALL
				SAMPLE	100%	OVERALL	SAMPLE					
							ACT.	EXP.	ACT.	EXP.	(OVERALL)	
1	Aug. 6	82	16,726	0	16,726	100.0	-	-	-	-	16.3	0.0
2	Aug. 13	169	33,995	365	33,630	100.0	-	-	-	-	15.6	0.0
3	Aug. 20	171	34,224	1,203	30,195	91.7	29.9	-	14.3	-	10.9	0.5
4	Aug. 26	213	42,726	4,505	27,080	73.9	28.8	33.6	10.1	15.9	9.3	1.4
5	Sep. 2	260	49,988	7,307	21,997	58.6	26.1	31.6	8.3	15.0	7.2	2.1
6	Sep. 7	228	45,072	8,531	12,172	45.9	25.9	29.6	7.2	12.0	6.2	2.5
7	Sep. 16	282	58,268	11,714	16,133	47.8	27.8	28.7	11.3	12.8	6.5	2.4
8	Sep. 23	336	65,799	14,199	11,996	39.8	26.4	28.2	10.6	13.1	5.7	2.7
9	Sep. 30	333	66,596	16,395	7,561	36.0	27.8	32.5	12.2	18.2	6.0	3.3
10	Oct. 7	382	75,001	17,517	7,440	34.6	27.4	29.7	12.8	15.2	5.2	3.0
11	Oct. 14	254	52,912	15,536	1,861	32.9	30.4	32.4	15.6	18.4	5.2	3.4
12	Oct. 21	353	73,341	19,923	289	27.6	27.3	28.6	12.5	14.2	4.5	3.3
13	Oct. 28	317	66,282	16,184	48	24.4	24.4	28.7	9.5	15.1	4.5	3.4
14	Nov. 4	320	66,032	16,215	49	24.5	24.5	29.1	8.8	13.4	4.6	3.4
15	Nov. 10	192	38,746	11,529	84	29.8	29.6	34.9	16.0	19.9	5.4	3.8
16	Nov. 18	225	48,897	14,507	26	29.7	29.6	32.2	15.6	19.5	5.3	3.7
17	Nov. 25	237	47,648	11,676	97	24.5	24.4	30.0	10.3	16.6	4.7	3.6
18	Dec. 2	247	49,638	13,958	206	28.5	28.2	30.4	14.0	17.0	4.8	3.4
	TOTAL	4,601	931,966	201,960	187,590	41.8 (34.3)*	27.1	30.3	11.8	15.7	6.3 (5.4)*	2.8

B-096F
C.3.

Contrôle qualitatif de la saisie des données du recensement de l'agriculture de 1981

Rapport récapitulatif

Section du contrôle qualitatif
D.M.E.I.A.
janvier 1982

Recensement de l'agriculture de 1981
Contrôle qualitatif de la saisie des données

A. INTRODUCTION

Ce document résume l'application du contrôle qualitatif à l'opération de la saisie des données dans le recensement de l'agriculture de 1981.

Ce projet avait pour but de tenter d'atteindre ou de dépasser un degré de qualité préétabli à la saisie des données, et, en même temps, de réduire au minimum le temps et le coût de la vérification par une méthode de vérification d'un échantillon. La limite moyenne de la qualité des données de sortie (LMQDS) a été fixée à 4.0 %, ce qui veut simplement dire qu'au moins 96 % de tous les enregistrements traités seraient libres d'erreurs introduites pendant l'entrée des données, après l'application du procédé de contrôle qualitatif.

B. MÉTHODE DE CONTRÔLE QUALITATIF

1. DIRECTIVES

- Les directives détaillées concernant l'entrée des données ont été établies par la Sous-division des méthodes d'enquête-recensement (SDMER) en collaboration avec la Section de la saisie des données de la DTD.

- Les procédés de vérification ont été élaborés par la Section du contrôle qualitatif de la DMEIA.

2. FORMATION DES VÉRIFICATEURS

La formation du personnel affecté à la vérification d'un échantillon a été donnée par la Section du CQ dans la DTD, pendant le poste de soir (c'est-à-dire de 16 h à 19 h 30).

Le calendrier de formation était le suivant :

CALENDRIER DE FORMATION DES VÉRIFICATEURS

DATE	NOMBRE DE VÉRIFICATEURS
13 juillet	4
14 juillet	4
20 juillet	4
21 juillet	4
4 août	4
6 août	3 + 1 opérateur principal
11 août	4 + 1 surveillant
13 août	(en anglais) \ Séance avec tous les opérateurs
14 août	(en français)) pour discuter des problèmes
18 août	4 + 1 surveillant
25 août	4 + 1 opérateur principal
27 août	4 + 1 opérateur principal
1er septembre	Séance avec tous les vérificateurs pour discuter des résultats et des problèmes
8 septembre	Séance avec tous les vérificateurs pour discuter des résultats et des problèmes
17 septembre	Séance avec tous les vérificateurs pour discuter des résultats et des problèmes
29 septembre	Séance avec tous les vérificateurs pour discuter des résultats et des problèmes
5 novembre	Séance avec tous les vérificateurs pour discuter des résultats et des problèmes

Dix cours de formation d'environ 3 heures ont été donnés, chacun comportant un exposé, des dicussions et un examen écrit.

Au total, 36 vérificateurs, 3 opérateurs principaux et 2 surveillants ont été formés. Dans l'ensemble, les cours de formation se sont bien déroulés, sauf dans le cas de 3 vérificateurs qui ont éprouvé quelques difficultés et ont dû suivre un deuxième cours.

Sept séances de révision supplémentaires ont eu lieu en présence de tous les opérateurs pour discuter des problèmes qui ont pu se poser et pour s'assurer que les directives étaient bien comprises.

3. PROGRESSION DES PRÉPOSÉS À L'ENTRÉE DES DONNÉES

Au début, le travail des opérateurs affectés à l'entrée des données a fait l'objet d'une vérification globale jusqu'à ce que les préposés aient traité un nombre d'enregistrements suffisant pour permettre une estimation sûre de leur taux d'erreur.

Plusieurs problèmes se sont posés au début du dépouillement, de sorte que la vérification d'un échantillon du travail des préposés à l'entrée des données a pris beaucoup plus de temps que prévu. Dans le but d'aider les préposés à progresser le plus vite possible à cette étape, on a calculé leur taux d'erreur tous les jours. Lorsque le taux d'erreur d'un préposé se stabilisait à 8 % ou moins, on lui attribuait un plan d'échantillonnage. Voici un calendrier indiquant le nombre de ces opérateurs qui étaient qualifiés pour la vérification d'un échantillon à chaque semaine de traitement des données.

CALENDRIER DES OPÉRATEURS QUALIFIÉS

SEMAINE	DATE	NOMBRE D'OPÉRATEURS QUALIFIÉS	TOTAL DES OPÉRATEURS QUALIFIÉS	SEMAINE	DATE	NOMBRE D'OPÉRATEURS QUALIFIÉS	TOTAL DES OPÉRATEURS QUALIFIÉS
1	5 août	2		6	4 sept	2	
1	6 août	1		6	8 sept	2	33
-	-	-	4	7	14 sept	1	33
3	17 août	1		8	15 sept	5	
3	20 août	3		8	22 sept	1	37
3	21 août	1	4	9	28 sept	1	35
4	24 août	2		10	2 oct	1	38
4	25 août	5		11	9 oct	3	42
4	26 août	2	19	12			40
5	31 août	4		13			40
5	1er sept	1		14			38
5	2 sept	1		15			36
5	3 sept	4	24	16			35
				17			33
				18			33

En tout, 43 opérateurs se sont qualifiés pour la vérification d'un échantillon. La plupart des premiers opérateurs qui ont progressé à cette étape ont suivi des plans de 7.0 % et de 8.0 % qu'on a par la suite modifiés à chaque semaine (généralement en diminuant ce pourcentage) en fonction de l'aptitude au traitement des opérateurs.

4. TESTS DE VÉRIFICATION

L'application de la vérification d'un échantillon comporte d'habitude un examen du rendement de chaque vérificateur au moyen de tests de vérification. Ces tests visent à assurer que les directives sont bien comprises et appliquées. On a informé les vérificateurs que leur rendement serait évalué à l'aide de ces tests mais, à cause des nombreux problèmes qui se sont présentés au début du dépouillement (difficultés non liées à la vérification) et en raison des contraintes

imposées par les calendriers d'exécution et du temps nécessaire pour appliquer ces tests, ce type d'évaluation n'a jamais eu lieu. On a simplement suivi de façon très générale le processus de vérification d'un échantillon en comparant les statistiques prévues aux données observées (voir le tableau à la section D).

5. RENSEIGNEMENTS TRANSMIS

a. Opérateurs

i) Tous les jours

Les préposés à l'entrée des données ont été informés chaque jour des erreurs importantes (et des erreurs d'interprétation possibles) ainsi que de leurs progrès.

On a donné également des renseignements aux vérificateurs sur l'exactitude avec laquelle ils inscrivaient les erreurs sur la carte de contrôle.

ii) Toutes les semaines

- Cellules contenant des erreurs :
  Au cours des 9 premières semaines, on a calculé le nombre de cellules contenant des erreurs, et la fréquence des erreurs pour chaque cellule a été transmise à chaque préposé.

  A partir de la 10^{e} semaine, les chiffres sur les cellules contenant des erreurs ont été communiqués seulement aux préposés soumis à une vérification globale.

- Taux d'erreur :
  Les préposés à l'entrée des données ont obtenu aussi leur taux d'erreur respectif. Ils ont manifesté beaucoup d'intérêt dans leurs rendements individuel et relatif.

- Renseignements personnels sur les préposés :
  Un imprimé sur lequel figuraient des données sur chaque lot traité a été distribué à tous les préposés à l'entrée des données.

b. Surveillants

i) Tous les jours

On a communiqué avec les surveillants tous les jours afin de les tenir au courant du rendement des opérateurs et de la marche des travaux.

ii) Toutes les semaines

A chaque semaine, les surveillants ont reçu une liste des préposés à l'entrée des données indiquant leur plan d'échantillonnage modifié.

6. TRAITEMENT HEBDOMADAIRE DES DONNÉES DU CONTRÔLE QUALITATIF

Les cartes de contrôle remplies ont été recueillies tous les jours, chaque semaine se terminant le mercredi. Ces cartes étaient codées et, le jeudi matin, elles étaient présentées à la DTD pour être introduites dans un ordinateur, ce qui se faisait d'habitude le jour même. Nos programmes de

totalisation des données sur la vérification et le contrôle qualitatif étaient ensuite appliqués à ces renseignements. Les résultats étaient analysés et, au plus tard le vendredi, ils étaient transmis aux opérateurs et aux surveillants.

Ce programme hebdomadaire a été mis en train la 4^{e} semaine de traitement (le vendredi 28 août) et retenu jusqu'à la fin du recensement (le vendredi 27 novembre).

C. APPLICATION DU CONTRÔLE QUALITATIF

Le contrôle qualitatif a été mis en application le 3 août, mais de nombreuses difficultés se sont présentées au cours du premier mois de traitement, ce qui a retardé la progression des préposés à l'entrée des données vers le stade de la vérification d'un échantillon.

En général, les préposés ont trouvé l'opération d'entrée des données assez complexe et il semble que cela a causé beaucoup de confusion et de frustration. Des séances de discussion ont été organisées pour les surveillants, les opérateurs principaux, les vérificateurs et les préposés à l'entrée des données qui ont signalé les problèmes suivants :

- Documents
  - entrées ambiguës, multiples et illisibles;
  - confusion causée par les accolades ({), le signe positif (+) et le signe négatif (-) inscrits à côté des données;
  - entrées et notes inscrites en différentes couleurs d'encre et de mine, etc.

- Directives concernant l'entrée des données
  . certaines parties portent à confusion;
  . elles sont généralement assez longues et comportent un grand nombre de mesures relatives aux exceptions.

Ces ambiguïtés et cette confusion ont provoqué de nombreux désaccords entre les vérificateurs et les préposés à l'entrée des données, ce qui par la suite a entraîné des taux d'erreur bien au-dessus de la moyenne. On a alors reconnu que ces taux représentaient davantage des taux de désaccord que des taux d'erreur, étant donné que, dans un bon nombre de cas, le vérificateur avait remplacé de bonnes entrées par une inscription inexacte.

Ces problèmes ont fait l'objet de discussions avec tous les membres de l'équipe de travail et on a appliqué les mesures suivantes :

- La Sous-division du recensement de l'agriculture a augmenté l'ampleur du contrôle manuel à l'étape de l'examen préliminaire, afin de rendre claires bon nombre des entrées les plus difficiles.

- La Division du traitement des données a simplifié les instructions relatives à l'entrée des données.

- La Section du contrôle qualitatif a continué de fournir des renseignements tous les jours aux préposés à l'entrée des données et de surveiller de près les taux d'erreur. Une formation supplémentaire a également été offerte (par de nouvelles séances de révision) à tous les préposés dans le but de réduire le nombre de désaccords entre les préposés et les vérificateurs.

Grâce à ces mesures, l'opération de saisie des données du recensement de l'agriculture a commencé à se dérouler de façon beaucoup plus régulière (à partir de septembre). D'une semaine à l'autre, les taux globaux de vérification et d'erreur ont diminué progressivement vers des niveaux acceptables (voir le tableau de la section suivante).

D. RÉSULTATS ET ANALYSE

Les chiffres suivants donnent les principaux résultats du contrôle qualitatif à l'étape du traitement :

Nombre total d'enregistrements introduits	931,966
Nombre total d'enregistrements vérifiés	201,960
Taux global de vérification	41.8 %
Taux de vérification d'échantillons	27.1 %
Taux prévu de vérification d'échantillons	30.3 %
Niveau moyen de qualité des données d'entrée (taux d'erreur)	6.3 %
Niveau moyen de qualité des données de sortie (NMQDS)	2.8 %

Lors du test du volume il était prévu que le taux d'erreur global tomberait à une valeur entre 4 % et 5 %. Or, le taux d'erreur mesuré pour l'ensemble des préposés au cours du traitement s'élevait à 6.3 %. Mais, tel qu'il a déjà été noté, les nombreux problèmes imprévus qui sont survenus au premier mois du traitement ont dans une mesure considérable aidé à atteindre ce taux élevé. Il est important de souligner que, après cette période (c'est-à-dire après la décision de

faire d'autres analyses préliminaires, d'offrir plus de formation et de simplifier les instructions relatives à l'entrée des données), le taux d'erreur global est descendu à 5.4 %. En outre, le taux d'erreur de l'ensemble des préposés qui avaient progressé à l'étape de la vérification d'un échantillon atteignait 4.8 % en moyenne.

A cause des taux d'erreurs plus élevés que prévu au début, il a fallu soumettre un grand nombre d'opérateurs à la vérification globale initiale, ce qui a entraîné un taux global de vérification de 41.8 %. Mais là encore, ce taux est tombé à 34.3 %, après le premier mois de traitement.

Le taux de vérification d'échantillons a été évalué à 27.1 %, ce qui est raisonnablement bon en comparaison avec le 30.3 % qui avait été prévu (c'est-à-dire le résultat d'un calcul théorique basé sur les probabilités d'acceptation prévues). On a évalué le taux de rejet à 11.8 %, valeur un peu en-dessous du 15.7 % qui avait été prévu au moyen d'un calcul fondé sur le degré de qualité observé à l'entrée. Ce résultat indique probablement qu'un nombre légèrement insuffisant de rejets ont été effectués pendant l'application du contrôle qualitatif.

Le niveau moyen de qualité des données de sortie a été évalué à 2.8 %, ce qui est aussi raisonnablement bon en comparaison avec la LMQDS de 4.0 %. Le tableau suivant présente un résumé hebdomadaire des données traitées.

RECENSEMENT DE L'AGRICULTURE DE 1981

RÉSUMÉ HEBDOMADAIRE DU CQ

LMQDS = 4 %

SE-MAINE	DATE TERMINÉE	NOMBRE DE LOTS	NOMBRE D'ENREGIS-TREMENTS INTRODUITS	NOMBRE D'ENREGIS-TREMENT VÉRIFIÉS		TAUX DE VÉRIFICATION			TAUX DE REJET		TAUX D'ERREUR	NMQDS GLOBAL
				VÉR. D'UN ÉCHANT.	VÉR. GLOBALE	VÉR. GLOBALE	VÉR. D'UN ÉCHANT.					
							OBS.	PRÉVU	OBS.	PRÉVU	(GLOBAL)	
1	6 août	82	16,726	0	16,726	100.0	-	-	-	-	16.3	0.0
2	13 août	169	33,995	365	33,630	100.0	-	-	-	-	15.6	0.0
3	20 août	171	34,224	1.203	30,195	91.7	29.9	-	14.3	-	13.9	0.5
4	26 août	213	42,726	4,505	27,080	73.9	28.8	33.6	10.1	15.9	9.3	1.4
5	2 sept.	260	49,988	7,307	21,997	58.6	26.1	31.6	8.3	15.6	7.2	1.1
6	7 sept.	228	45,072	8,531	12,172	45.9	25.9	29.6	7.2	12.0	6.2	2.5
7	16 sept.	282	58,268	11,714	16,133	47.8	27.8	28.7	11.3	12.8	6.5	2.4
8	23 sept.	336	65,799	14,199	11,996	39.8	26.4	28.2	10.6	13.1	5.7	2.7
9	30 sept.	333	66,596	16,395	7,561	36.0	27.8	32.5	12.2	18.3	6.6	3.3
10	7 oct.	382	75,001	17,517	7,440	34.6	27.4	29.7	12.8	15.2	7.3	3.0
11	14 oct.	254	52,912	15,536	1,861	32.9	30.4	32.4	15.6	18.3	6.2	3.4
12	21 oct.	353	73,341	19,923	289	27.6	27.3	28.6	12.5	14.3	6.5	3.3
13	28 oct.	317	66,282	16,184	48	24.4	24.4	28.7	9.5	15.1	6.5	3.4
14	4 nov.	320	66,032	16,215	49	24.5	24.5	29.1	8.8	13.4	5.6	3.4
15	10 nov.	192	38,746	11,529	84	29.8	29.6	34.9	16.0	19.9	5.4	3.8
16	18 nov.	225	48,897	14,507	26	29.7	29.6	32.2	15.6	19.5	5.3	3.7
17	25 nov.	237	47,648	11,676	97	24.5	24.4	30.0	10.3	16.6	4.7	3.6
18	2 dec.	247	49,638	13,958	206	28.5	28.2	30.4	14.0	17.0	5.8	3.4
	TOTAL	4,601	931,966	201,960	187,590	41.8 (34.3)*	27.1	30.3	11.8	15.7	6.3 (5.4)*	2.8

Notes sur ce tableau

1 - Semaine 2 : La vérification d'un échantillon a été appliquée à deux lots seulement et les deux ont été rejetés.

2 - Semaine 2-3 : Le nombre de lots soumis à la vérification d'un échantillon est trop petit pour permettre le calcul des taux prévus.

3 - Pour certaines semaines, on a constaté que les taux de vérification et de rejet prévus étaient plus élevés que les taux observés, suffisamment pour qu'on soupçonne une certaine tendance à rejeter moins de lots que normalement. Plusieurs séances ont donc eu lieu avec les vérificateurs pour discuter de ce problème et, en même temps, revoir les directives concernant la vérification d'un échantillon.

4 - Les nombres marqués d'un astérisque (*) dans les totaux correspondent à des taux calculés après le premier mois de traitement.

E. PIÈCES JOINTES

Les pièces suivantes sont jointes à ce document:

1 - Un graphique de la qualité à l'entrée et de la qualité à la sortie en fonction de la semaine de traitement.

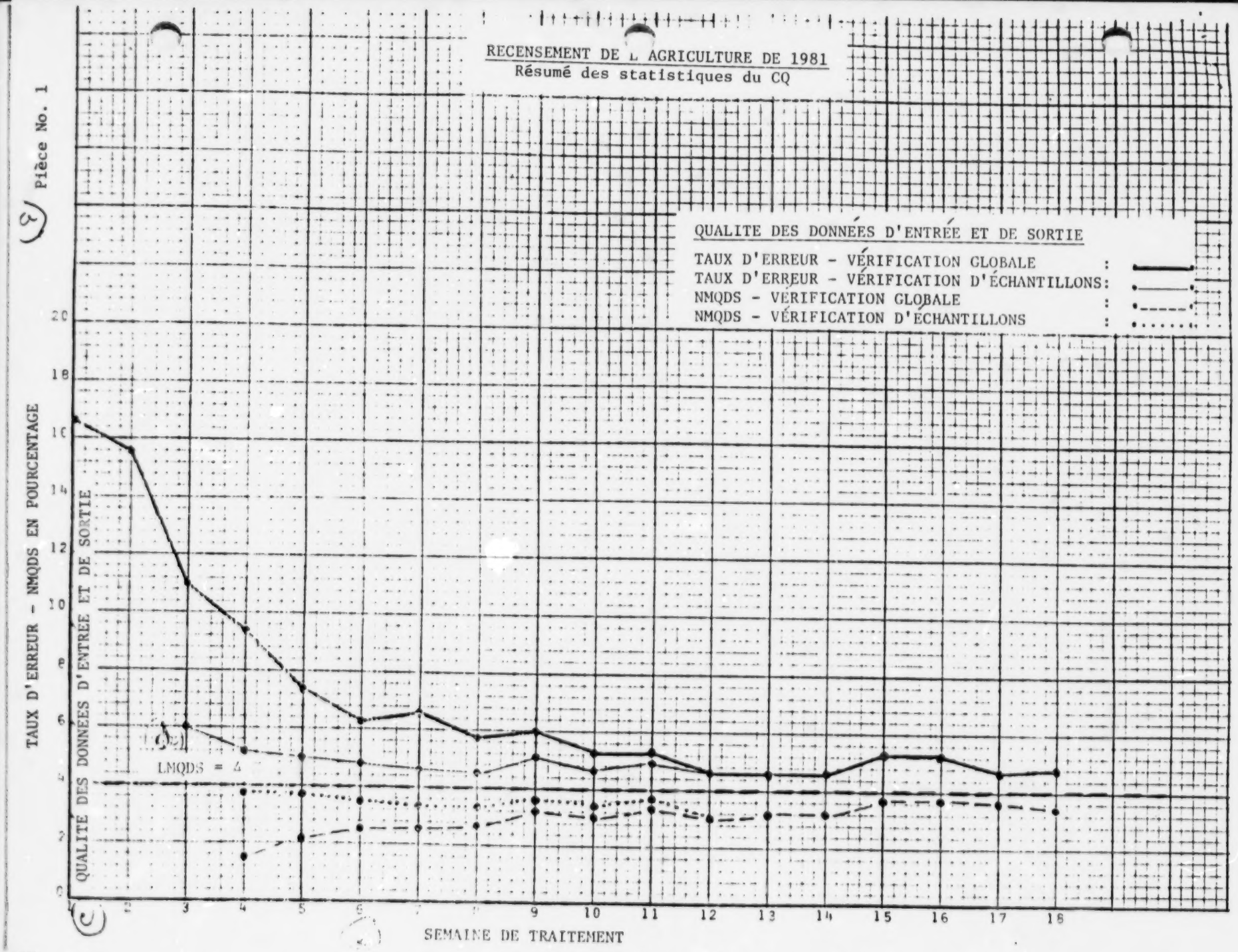
2 - Un graphique des taux de vérification et de rejet en fonction de la semaine de traitement.

3 - Un graphique du taux de vérification cumulé en fonction de la semaine de traitement.

4 - Une liste détaillée des cellules contenant une erreur par numéro de cellule, la fréquence des erreurs et leur distribution exprimée en pourcentage de l'erreur totale.

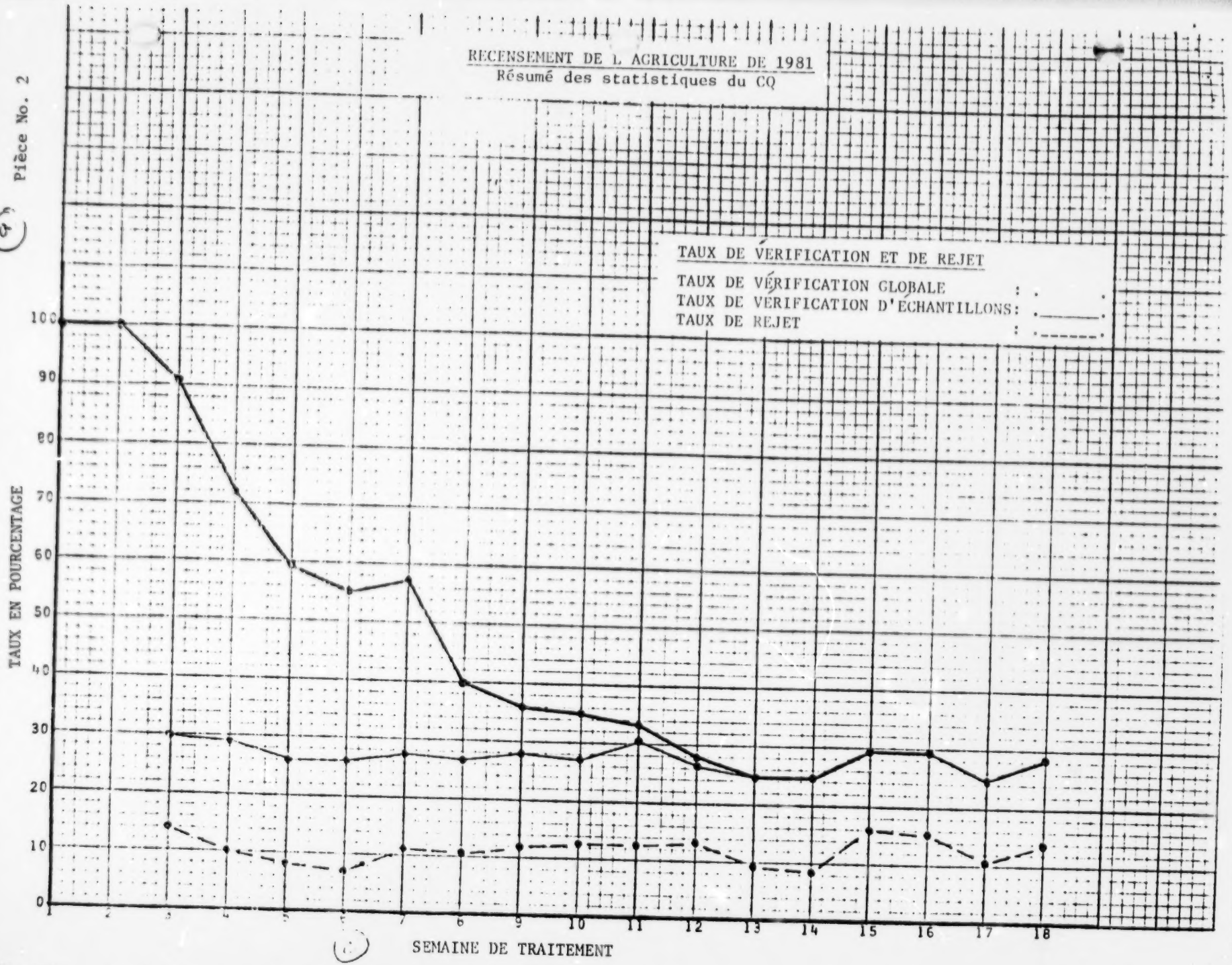
Pièce No. 1

RECENSEMENT DE L AGRICULTURE DE 1981
Résumé des statistiques du CQ


QUALITE DES DONNÉES D'ENTRÉE ET DE SORTIE
TAUX D'ERREUR - VÉRIFICATION GLOBALE :
TAUX D'ERREUR - VÉRIFICATION D'ÉCHANTILLONS:
NMQDS - VÉRIFICATION GLOBALE :
NMQDS - VÉRIFICATION D'ÉCHANTILLONS :
TAUX D'ERREUR - NMQDS EN POURCENTAGE
QUALITE DES DONNEES D'ENTREE ET DE SORTIE
LMQDS = 4
SEMAINE DE TRAITEMENT

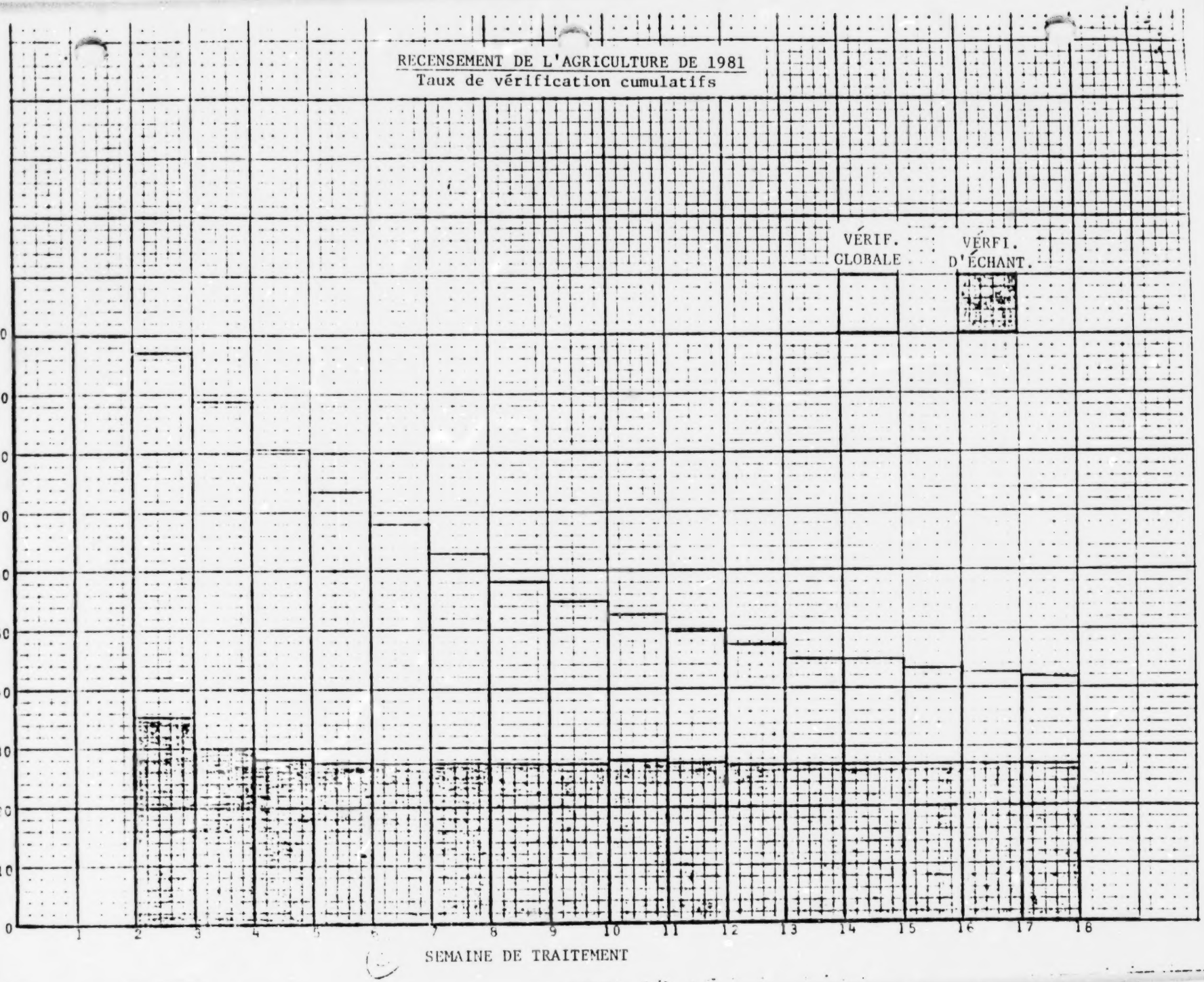
Pièce No. 2

RECENSEMENT DE L AGRICULTURE DE 1981
Résumé des statistiques du CQ



Pièce No. 3

RECENSEMENT DE L'AGRICULTURE DE 1981
Taux de vérification cumulatifs



Pièce No. 4

RECENSEMENT DE L'AGRICULTURE DE 1981

CELLULES CONTENANT DES ERREURS

# DE CELL	FREQ	DISTR. %	# DE LA CELL	FREQ	DISTR. %	# DE LA CELL	FREQ	DISTR. %	# DE LA CELL	FREQ	DISTR. %	# DE LA CELL	FREQ	DISTR. %
001	39	0.1	061	6	0.02	109	9	0.02	157	93	0.2	205	175	0.5
002	34	0.09	062	6	0.02	110	9	0.02	158	159	0.4	206	65	0.2
003	69	0.2	063	3	0.008	111	15	0.04	159	299	0.8	207	114	0.3
004	730	2.0	064	-	0	112	3	0.008	160	422	1.1	208	209	0.5
005	568	1.5	065	1	0.003	113	10	0.03	161	556	1.4	209	277	0.7
006	47	0.1	066	4	0.01	114	36	0.09	162	459	1.2	210	28	0.07
007	132	0.3	067	3	0.008	115	20	0.05	163	49	0.1	211	49	0.1
008	306	0.8	068	1	0.003	116	11	0.03	164	70	0.2	212	74	0.2
009	137	0.4	069	29	0.07	117	14	0.04	165	99	0.3	213	100	0.3
010	62	0.2	070	41	0.1	118	6	0.02	166	103	0.3	214	741	1.9
011	85	0.2	071	57	0.2	119	34	0.09	167	67	0.1	215	1061	2.8
024	46	0.1	072	225	0.6	120	8	0.02	168	122	0.3	216	111	0.3
025	174	0.5	073	15	0.04	121	49	0.1	169	32	0.08	217	229	0.6
026	189	0.5	074	43	0.1	122	97	0.2	170	23	0.06	218	463	1.2
027	128	0.3	075	68	0.2	123	75	0.2	171	27	0.07	219	169	0.4
028	2	0.005	076	10	0.03	124	7	0.02	172	151	0.4	220	285	0.7
029	814	2.1	077	37	0.1	125	748	1.9	173	45	0.1	221	665	1.7
030	872	2.3	078	64	0.2	126	440	1.1	174	39	0.1	222	642	1.7
031	770	2.0	079	14	0.04	127	250	0.7	175	2	0.005	223	679	1.8
032	421	1.1	080	35	0.09	128	695	1.8	176	4	0.01	224	519	1.4
033	1089	2.8	081	71	0.2	129	462	1.2	177	7	0.02	225	362	0.9
034	155	0.4	082	4	0.01	130	437	1.1	178	40	0.1	226	731	1.9
035	43	0.1	083	13	0.03	131	815	2.1	179	36	0.09	227	710	1.9
036	23	0.06	084	27	0.07	132	72	0.2	180	165	0.4	228	570	1.5
037	61	0.2	085	8	0.02	133	11	0.03	181	83	0.3	229	592	1.5
038	190	0.5	086	13	0.03	134	5	0.01	182	146	0.4	230	151	0.4
039	166	0.4	087	50	0.1	135	17	0.04	183	214	0.6	231	92	0.2
040	76	0.2	088	7	0.02	136	45	0.1	184	68	0.2	232	835	2.2
041	131	0.3	089	23	0.06	137	3	0.008	185	70	0.2	233	87	0.2
042	22	0.06	090	32	0.08	138	28	0.07	186	117	0.3	234	245	0.6
043	11	0.03	091	135	0.4	139	18	0.05	187	37	0.1	AJOUT SUPPR	605	1.6
044	8	0.02	092	66	0.2	140	333	0.9	188	368	1.0			
045	19	0.05	093	97	0.3	141	441	1.1	189	680	1.8	# DE QUEST	65	0.2
046	49	0.1	094	49	0.1	142	161	0.4	190	320	0.8			
047	4	0.01	095	65	0.2	144	159	0.4	191	727	1.9	# D'ENV	90	0.2
048	28	0.07	096	31	0.08	144	92	0.2	192	331	0.9			
049	2	0.005	097	30	0.08	145	109	0.3	193	470	1.2	TOT	38,376	100.0
050	90	0.2	098	49	0.1	146	269	0.7	194	251	0.7			
051	216	0.6	099	32	0.1	147	135	0.4	195	429	1.1			
052	260	0.7	100	35	0.09	148	190	0.4	196	101	0.3			
053	60	0.2	101	82	0.2	149	47	0.1	197	231	0.6			
054	47	0.1	102	54	0.1	150	35	0.09	198	55	0.1			
055	28	0.07	103	30	0.08	151	168	0.4	199	154	0.4			
056	19	0.05	104	44	0.1	152	165	0.4	200	113	0.3			
057	4	0.01	105	41	0.1	153	170	0.4	201	245	0.6			
058	9	0.02	106	155	0.4	154	238	0.6	202	59	0.2			
059	11	0.03	107	101	0.3	155	197	0.5	203	110	0.3			
060	-	0	108	7	0.02	156	172	0.4	204	108	0.3			

Ce tableau présente des chiffres sur le nombre de fois que chaque cellule a été enregistrée comme contenant une erreur sur la carte de contrôle.

La colonne de la distribution en pourcentage est calculée comme un pourcentage du nombre total d'erreurs repérées. Les fréquences relatives ne sont pas indiquées parce que la fréquence à l'intérieur de chaque cellule est inconnue.

POOR

QUALITY

ORIGINAL